Bilinear forms and their matrices

Joel Kamnitzer

March 11, 2011

0.1 Definitions

A *bilinear form* on a vector space V over a field $\mathbb{F}$ is a map

$$H : V \times V \to \mathbb{F}$$

such that

(i) $H(v_1 + v_2, w) = H(v_1, w) + H(v_2, w),$ for all $v_1, v_2, w \in V$

(ii) $H(v, w_1 + w_2) = H(v, w_1) + H(v, w_2),$ for all $v, w_1, w_2 \in V$

(iii) $H(av, w) = aH(v, w),$ for all $v, w \in V, a \in \mathbb{F}$

(iv) $H(v, aw) = aH(v, w),$ for all $v, w \in V, a \in \mathbb{F}$

A bilinear form H is called *symmetric* if $H(v, w) = H(w, v)$ for all $v, w \in V$.

A bilinear form H is called *skew-symmetric* if $H(v, w) = -H(w, v)$ for all $v, w \in V$.

A bilinear form H is called *non-degenerate* if for all $v \in V$, there exists $w \in V$, such that $H(w, v) \neq 0$.

A bilinear form H defines a map $H^\# : V \to V^*$ which takes w to the linear map $v \mapsto H(v, w)$. In other words, $H^\#(w)(v) = H(v, w)$.

Note that H is non-degenerate if and only if the map $H^\# : V \to V^*$ is injective. Since V and V^* are finite-dimensional vector spaces of the same dimension, this map is injective if and only if it is invertible.

0.2 Matrices of bilinear forms

If we take $V = \mathbb{F}^n$, then every $n \times n$ matrix A gives rise to a bilinear form by the formula

$$H_A(v, w) = v^t A w$$

Example 0.1. Take $V = \mathbb{R}^2$. Some nice examples of bilinear forms are the ones coming from the matrices:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Conversely, if V is any vector space and if $v_1, \dots, v_n$ is a basis for V, then we define the matrix $[H]_{v_1,\dots,v_n}$ for H with respect to this basis to be the matrix whose i, j entry is $H(v_i, v_j)$.

Proposition 0.2. *Take $V = \mathbb{F}^n$. The matrix for H_A with respect to the standard basis is A itself.*

Proof. By definition,

$$H_A(e_i, e_j) = e_i^t A e_j = A_{ij}.$$

□

Recall that if V is a vector space with basis $v_1, \dots, v_n$, then its dual space V^* has a dual basis $\alpha_1, \dots, \alpha_n$. The element α_j of the dual basis is defined as the unique linear map from V to $\mathbb{F}$ such that

$$\alpha_j(v_i) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise} \end{cases}$$

Proposition 0.3. *The matrix for H with respect to $v_1, \dots, v_n$ is the same as the matrix for $v_1, \dots, v_n$ and $\alpha_1, \dots, \alpha_n$ with respect to the map $H^\# : V \to V^*$.*

Proof. Let $A = [H^\#]_{v_1,\dots,v_n}^{\alpha_1,\dots,\alpha_n}$. Then

$$H^\# v_j = \sum_{k=1}^{n} A_{kj} \alpha_k$$

Hence, $H(v_i, v_j) = H^\#(v_j)(v_i) = A_{ij}$ as desired. □

From this proposition, we deduce the following corollary.

Corollary 0.4. *H is non-degenerate if and only if the matrix $[H]_{v_1,\dots,v_n}$ is invertible.*

It is interesting to see how the matrix for a bilinear form changes when we changes the basis.

Theorem 0.5. *Let V be a vector space with two bases $v_1, \dots, v_n$ and $w_1, \dots, w_n$. Let Q be the change of basis matrix. Let H be a bilinear form on V.*

Then

$$Q^t [H]_{v_1,\dots,v_n} Q = [H]_{w_1,\dots,w_n}$$

Proof. Choosing the basis $v_1, \dots, v_n$ means that we can consider the case where $V = \mathbb{F}^n$, and $v_1, \dots, v_n$ denotes the standard basis. Then $w_1, \dots, w_n$ are the columns of Q and $w_i = Qv_i$.

Let $A = [H]_{v_1,\dots,v_n}$.

So we have

$$H(w_i, w_j) = w_i^t A w_j = (Qv_i)^t A Q v_j = v_i^t Q^t A Q v_j$$

as desired. □

You can think of this operation $A \mapsto Q^t A Q$ as simultaneous row and column operations.

Example 0.6. Consider

$$A = \begin{bmatrix} 0 & 4 \\ 4 & 2 \end{bmatrix}$$

After doing simultaneous row and column operations we reach

$$Q^t A Q = \begin{bmatrix} -8 & 0 \\ 0 & 2 \end{bmatrix}$$

The new basis is $(1, -2), (0, 1)$.

0.3 Isotropic vectors and perp spaces

A vector v is called *isotropic* if $H(v, v) = 0$.

If H is skew-symmetric, then $H(v, v) = -H(v, v)$, so every vector is isotropic.

Let H be a non-degenerate bilinear form on a vector space V and let $W \subset V$ be a subspace. We define the *perp space* to W as

$$W^\perp = \{v \in V : H(w, v) = 0 \text{ for all } w \in W\}$$

Notice that $W^\perp$ may intersect W. For example if W is the span of a vector v, then $W \subset W^\perp$ if and only if v is isotropic.

Example 0.7. If we take $\mathbb{R}^2$ with the bilinear form $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, then $(1, 1)$ is an isotropic vector and $span(1, 1)^\perp = span(1, 1)$.

So in general, V is not the direct sum of W and $W^\perp$. However, we have the following result which says that they have complementary dimension.

Proposition 0.8. $\dim W^\perp = \dim V - \dim W$

Proof. We have defined $H^\# : V \to V^*$. The inclusion of $W \subset V$ gives us a surjective linear map $\pi : V^* \to W^*$, and so by composition we get $T = \pi \circ H^\# : V \to W^*$. This map T is surjective since $H^\#$ is an isomorphism. Thus

$$\dim \text{null}(T) = \dim V - \dim W^* = \dim V - \dim W$$

Checking through the definitions, we see that

$$v \in \text{null}(T) \text{ if and only if } H^\#(v)(w) = 0 \text{ for all } w \in W$$

Since $H^\#(v)(w) = H(w, v)$, this shows that $v \in \text{null}(T)$ if and only if $v \in W^\perp$. Thus $W^\perp = \text{null}(T)$ and so the result follows. $\square$

Symmetric bilinear forms

Joel Kamnitzer

March 14, 2011

1 Symmetric bilinear forms

We will now assume that the characteristic of our field is not 2 (so $1+1 \neq 0$).

1.1 Quadratic forms

Let H be a symmetric bilinear form on a vector space V. Then H gives us a function $Q : V \to \mathbb{F}$ defined by $Q(v) = H(v, v)$. Q is called a quadratic form. We can recover H from Q via the equation

$$H(v, w) = \frac{1}{2}(Q(v+w) - Q(v) - Q(w))$$

Quadratic forms are actually quite familiar objects.

Proposition 1.1. *Let $V = \mathbb{F}^n$. Let Q be a quadratic form on $\mathbb{F}^n$. Then $Q(x_1, \ldots, x_n)$ is a polynomial in n variables where each term has degree 2. Conversely, every such polynomial is a quadratic form.*

Proof. Let Q be a quadratic form. Then

$$Q(x_1, \ldots, x_n) = [x_1 \cdots x_n] A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

for some symmetric matrix A.

Expanding this out, we see that

$$Q(x_1, \ldots, x_n) = \sum_{1 \leq i,j \leq n} A_{ij} x_i x_j$$

and so it is a polynomial with each term of degree 2. Conversely, any polynomial of degree 2 can be written in this form. □

Example 1.2. Consider the polynomial $x^2 + 4xy + 3y^2$. This the quadratic form coming from the bilinear form H_A defined by the matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$.

We can use this knowledge to understand the graph of solutions to $x^2 + 4xy + 3y^2 = 1$. Note that H_A has a diagonal matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ with respect to the basis $(1, 0), (-2, 1)$. This shows that $Q(a(1, 0) + b(-2, 1)) = a^2 - b^2$. Thus the solutions of $x^2+4xy+3y^2 = 1$ are obtained from the solutions to $a^2-b^2 = 1$ by a linear transformation. Thus the graph is a hyperbola.

1.2 Diagonalization

As we saw before, the bilinear form is symmetric if and only if it is represented by a symmetric matrix. We now will consider the problem of finding a basis for which the matrix is diagonal. We say that a bilinear form is *diagonalizable* if there exists a basis for V for which H is represented by a diagonal matrix.

Lemma 1.3. *Let H be a non-trivial bilinear form on a vector space V. Then there exists $v \in V$ such that $H(v, v) \neq 0$.*

Proof. There exist $u, w \in V$ such that $H(u, w) \neq 0$. If $H(u, u) \neq 0$ or $H(w, w) \neq 0$, then we are done. So we assume that both u, w are isotropic. Let $v = u + w$. Then $H(v, v) = 2H(u, w) \neq 0$. □

Theorem 1.4. *Let H be a symmetric bilinear form on a vector space V. Then H is diagonalizable.*

This means that there exists a basis $v_1, \dots, v_n$ for V for which $[H]_{v_1,\dots,v_n}$ is diagonal, or equivalently that $H(v_i, v_j) = 0$ if $i \neq j$.

Proof. We proceed by induction on the dimension of the vector space V. The base case is $\dim V = 0$, which is immediate. Assume the result holds for all bilinear forms on vector spaces of dimension $n - 1$ and let V be a vector space of dimension n.

If $H = 0$, then we are already done. Assume $H \neq 0$, then by the Lemma we get $v \in V$ such that $H(v, v) \neq 0$.

Let $W = \operatorname{span}(v)^\perp$. Since v is not isotropic, $W \oplus \operatorname{span}(v) = V$. Since $\dim W = n - 1$, the result holds for W. So pick a basis $v_1, \dots, v_{n-1}$ for W for which H_W is diagonal and then extend to a basis $v_1, \dots, v_{n-1}, v$ for V. Since $v_i \in W$, $H(v, v_i) = 0$ for $i = 1, \dots, n-1$. Thus the matrix for H is diagonal. □

1.3 Diagonalization in the real case

For this section we will mostly work with real vector spaces. Recall that a symmetric bilinear form H on a real vector space V is called *positive definite* if $H(v, v) > 0$ for all $v \in V$, $v \neq 0$. A postive-definite symmetric bilinear form is the same thing as an inner product on V.

Theorem 1.5. *Let H be a symmetric bilinear form on a real vector space V. There exists a basis $v_1, \dots, v_n$ for V such that $[H]_{v_1,\dots,v_n}$ is diagonal and all the entries are $1, -1,$ or 0.*

We have already seen a special case of this theorem. Recall that if H is an inner product, then there is an orthonormal basis for H. This is the same as a basis for which the matrix for H consists of just 1s on the diagonal.

Proof. By the previous theorem, we can find a basis $w_1, \dots, w_n$ for V such that $H(w_i, w_j) = 0$ for $i \neq j$. Let $a_i = H(w_i, w_i)$ for $i = 1, \dots, n$. Define

$$v_i = \begin{cases} \frac{1}{\sqrt{a_i}} w_i, \text{ if } a_i > 0 \\ \frac{1}{\sqrt{-a_i}} w_i, \text{ if } a_i < 0 \\ w_i, \text{ if } a_i = 0 \end{cases} \tag{1}$$

Then $H(v_i, v_i)$ is either $1, -1$, or 0 depending on the three cases above. Also $H(v_i, v_j) = 0$ for $i \neq j$ and so we have found the desired basis. □

Corollary 1.6. *Let Q be a quadratic form on a vector space V. There exists a basis $v_1, \dots, v_n$ for V such that the quadratic form is given by*

$$Q(x_1 v_1 + \cdots + x_n v_n) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$$

Proof. Let H be a associated bilinear form. Pick a basis $v_1, \dots, v_n$ as in the theorem, ordered so that the diagonal entries in the matrix are 1s then -1s, then 0s. The result follows. □

Given a symmetric bilinear form H on a real vector space V, pick a basis $v_1, \dots, v_n$ for V as above. Let p be the number of 1s and q be the number of -1s in the diagonal entries of the matrix. The following result is known (for some reason) as "Sylveter's Law of Inertia".

Theorem 1.7. *The numbers p, q depend only on the bilinear form. (They do not depend on the choice of basis $v_1, \dots, v_n$.)*

To prove this result, we will begin with the following discussion which applies to symmetric bilinear forms over any field. Given a symmetric bilinear form H, we define its radical (sometimes also called kernel) to be

$$\mathrm{rad}(H) = \{w \in V : H(v, w) = 0 \text{ for all } v \in V\}$$

In other words, $\mathrm{rad}(H) = V^{\perp}$. Another way of thinking about this is to say that $\mathrm{rad}(H) = \mathrm{null}(H^{\#})$.

Lemma 1.8. *Let H be a symmetric bilinear form on a vector space V. Let $v_1, \dots, v_n$ be a basis for V and let $A = [H]_{v_1, \dots, v_n}$. Then*

$$\dim \mathrm{rad}(H) = \dim V - \mathrm{rank}(A)$$

Proof. Recall that A is actually the matrix for the linear map $H^{\#}$. Hence $\mathrm{rank}(A) = \mathrm{rank}(H^{\#})$. So the result follows by the rank-nullity theorem for $H^{\#}$. □

Proof of Theorem 1.7. The lemma shows us that $p+q$ is an invariant of H. So it suffices to show that p is independent of the basis.

Let

$$\tilde{p} = \max\left(\dim W : W \text{ is a subspace of } V \text{ and } H|_W \text{ is positive definite}\right)$$

Clearly, $\tilde{p}$ is independent of the basis. We claim that $p = \tilde{p}$.

Assume that our basis $v_1, \dots, v_n$ is ordered so that

$$\begin{aligned} &H(v_i, v_i) = 1 \text{ for } i = 1, \dots p, \\ &H(v_i, v_i) = -1 \text{ for } i = p+1, \dots, p+q, \text{ and} \\ &H(v_i, v_i) = 0 \text{ for } i = p+q+1, \dots, n \end{aligned}$$

Let $W = \operatorname{span}(v_1, \dots, v_p)$. Then $\dim W = p$ and so $p \le \tilde{p}$.

To see that $\tilde{p} \le p$, let $\tilde{W}$ be a subspace of V such that $H|_{\tilde{W}}$ is positive definite and $\dim \tilde{W} = \tilde{p}$.

We claim that $\tilde{W} \cap \operatorname{span}(v_{p+1}, \dots, v_n) = 0$. Let $v \in \tilde{W} \cap \operatorname{span}(v_{p+1}, \dots, v_n)$, $v \neq 0$. Then $H(v, v) > 0$ by the definition of $\tilde{W}$. On the other hand, if $v \in \operatorname{span}(v_{p+1}, \dots, v_n)$, then

$$v = x_{p+1} v_{p+1} + \cdots + x_n v_n$$

and so $H(v, v) = -x_{p+1}^2 - \cdots - x_{p+q}^2 \le 0$. We get a contradiction. Hence $\tilde{W} \cap \operatorname{span}(v_{p+1}, \dots, v_n) = 0$.

This implies that

$$\dim \tilde{W} + \dim \operatorname{span}(v_{p+1}, \dots, v_n) \le n$$

and so $\tilde{p} \le n - (n-p) = p$ as desired. □

The pair (p, q) is called the signature of the bilinear form H. (Some authors use $p - q$ for the signature.)

Example 1.9. Consider the binear form on $\mathbb{R}^2$ given by the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. It has signature $(1, -1)$.

Example 1.10. In special relativity, symmetric bilinear forms of signature $(3, 1)$ are used.

In the complex case, the theory simplifies considerably.

Theorem 1.11. *Let H be a symmetric bilinear form on a complex vector space V. Then there exists a basis $v_1, \dots, v_n$ for V for which $[H]_{v_1, \dots, v_n}$ is a diagonal matrix with only 1s or 0s on the diagonal. The number of 0s is the dimension of the radical of H.*

Proof. We follow the proof of Theorem 1.5. We start with a basis $w_1, \dots, w_n$ for which the matrix of H is diagonal. Then for each i with $H(w_i, w_i) \neq 0$, we choose a_i such that $a_i^2 = \frac{1}{H(w_i, w_i)}$. Such a_i exists, since we are working with complex numbers. Then we set $v_i = a_i w_i$ as before. □

Symplectic forms

Joel Kamnitzer

March 24, 2011

1 Symplectic forms

We assume that the characteristic of our field is not 2 (so $1 + 1 \neq 0$).

1.1 Definition and examples

Recall that a *skew-symmetric bilinear form* is a bilinear form Ω such that $\Omega(v, w) = -\Omega(w, v)$ for all $v, w \in V$. Note that if Ω is a skew-symmetric bilinear form, then $\Omega(v, v) = 0$ for all $v \in V$. In other words, every vector is isotropic.

A *symplectic form* is a non-degenerate skew-symmetric bilinear form. Recall that non-degenerate means that for all $v \in V$ such that $v \neq 0$, there exists $w \in V$ such that $\Omega(v, w) \neq 0$.

Example 1.1. Consider $V = \mathbb{F}^2$ and take the bilinear form given by the matrix

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Here is a more general example. Let W be a vector space. Define a vector space $V = W \oplus W^*$. We define a bilinear form on W by the rule

$$\Omega((v_1, \alpha_1), (v_2, \alpha_2)) = \alpha_2(v_1) - \alpha_1(v_2)$$

This is clearly skew-symmetric. It is also non-degenerate since if $(v_1, \alpha_1) \in V$ is non-zero, then either $v_1 \neq 0$ or $\alpha_1 \neq 0$. Assume that $v_1 \neq 0$. Then we can choose $\alpha_2 \in V^*$ such that $\alpha_2(v_1) \neq 0$ and so $\Omega((v_1, \alpha_1), (0, \alpha_2)) \neq 0$. So Ω is a symplectic form.

1.2 Symplectic bases

We cannot hope to diagonalize a symplectic form since every vector is isotropic. We will instead introduce a different goal.

Let V, Ω be a vector space with a symplectic form. Suppose that $\dim V = 2n$. A *symplectic basis* for V is a basis $q_1, \ldots, q_n, p_1, \ldots, p_n$ for V such that

$$\Omega(p_i, q_i) = 1$$
$$\Omega(q_i, p_i) = -1$$
$$\Omega(p_i, q_j) = 0 \text{ if } i \neq j$$
$$\Omega(p_i, p_j) = 0$$
$$\Omega(q_i, q_j) = 0$$

In other words the matrix for Ω with respect to this basis is the $2n \times 2n$ matrix

$$\begin{bmatrix} 0_n & -I_n \\ I_n & 0_n \end{bmatrix}$$

Theorem 1.2. *Let V be a vector space and let Ω be a symplectic form. Then $\dim V$ is even and there exists a symplectic basis for V.*

Proof. We proceed by induction on $\dim V$. The base cases are $\dim V = 0$ and $\dim V = 1$.

Now assume $\dim V = n$ and assume the result holds for all vector spaces of dimension $n - 2$. Let $q \in V$, $q \neq 0$. Since Ω is non-degenerate, there exists $p \in V$ such that $\Omega(p, q) \neq 0$. By scaling q, we can ensure that $\Omega(p, q) = 1$. Then $\Omega(q, p) = -1$ by skew symmetry.

Let $W = span(p, q)^{\perp}$. So

$$W = \{v \in V : \Omega(v, p) = 0 \text{ and } \Omega(v, q) = 0\}.$$

We claim that $W \cap span(p, q) = 0$. Let $v \in W \cap span(p, q)$. Then $v = ap + bq$ for some $a, b \in \mathbb{F}$. Since $v \in W$, we see that $\Omega(v, p) = 0$. But $\Omega(v, p) = -b$, so $b = 0$. Similarly, $\Omega(v, q) = 0$, which implies that $a = 0$. Hence $v = 0$.

Since Ω is non-degenerate, we know that $\dim W + \dim span(p, q) = \dim V$. Thus $W \oplus span(p, q) = V$.

To apply the inductive hypothesis, we need to check now that the restriction of Ω to W is a symplectic form. It is clearly skew-symmetric, so we just need to check that it is non-degenerate. To see this, pick $w \in W$, $w \neq 0$. Then there exists $v \in V$ such that $\Omega(w, v) \neq 0$. We can write $w = u + u'$, where $u \in W$ and $u' \in span(p, q)$. By the definition of $span(p, q)$, $\Omega(u', w) = 0$. Hence $\Omega(u, w) \neq 0$ and so the restriction of Ω to W is non-degenerate.

We now apply the inductive hypothesis to W. Note that $\dim W = \dim V - 2$. By the inductive hypothesis, $\dim W$ is even and we have a symplectic basis $q_1, \ldots, q_m, p_1, \ldots p_m$ where $2m = n - 2$. Hence $\dim V$ is even. We claim that $q_1, \ldots, q_m, q, p_1, \ldots, p_m, p$ is a symplectic basis for W. This follows from the definitions. $\square$

1.3 Lagrangians

Let V be a vector space of dimension $2n$ and Ω be a symplectic form on V. Recall that a subspace W of V is called *isotropic* if $\Omega(w_1, w_2) = 0$ for all $w_1, w_2 \in W$. This is equivalent to the condition that $W \subset W^\perp$. Since Ω is non-degenerate, $\dim W^\perp + \dim W = \dim V$. Hence the maximum possible dimension of an isotropic subspace is n. An isotropic subspace L of dimension n is called a *Lagrangian.*

Example 1.3. Let V be a 2 dimensional vector space. Then any 1-dimensional subspace is Lagrangian.

We can also produce Lagrangian subspaces from symplectic bases as follows.

Proposition 1.4. *Let $q_1, \ldots, q_n, p_1, \ldots, p_n$ be a symplectic basis for V. Then $span(q_1, \ldots, q_n), span(p_1, \ldots, p_n)$ are both Lagrangian subspaces of V.*

Proof. From the fact that $\Omega(q_i, q_j) = 0$ for all i, j, we see that $\Omega(v, w) = 0$ for all $v, w \in span(q_1, \ldots, q_n)$. Hence $span(q_1, \ldots, q_n)$ is isotropic. Since it has dimension n, it is Lagrangian. The result for $p_1, \ldots, p_n$ is similar. $\square$

Now suppose that $V = W \oplus W^*$ for some vector space W and we define a symplectic form Ω on W as above. Then it is easy to see that W and W^* are both Lagrangian subspaces of V.

Multilinear forms

Joel Kamnitzer

April 1, 2011

Assume that all fields are characteristic 0 (i.e. $1 + \cdots + 1 \neq 0$), for example $\mathbb{F} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$. Assume also that all vector spaces are finite dimensional.

1 Dual spaces

If V is a vector space, then $V^* = L(V, \mathbb{F})$ is defined to be the space of linear maps from V to $\mathbb{F}$.

If $v_1, \ldots, v_n$ is a basis for V, then we define $\alpha_i \in V^*$ for $i = 1, \ldots, n$, by setting

$$\alpha_i(v_j) = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$$

Proposition 1.1. *$\alpha_1, \ldots, \alpha_n$ forms a basis for V^* (called the dual basis).*

In particular, this shows that V and V^* are vector spaces of the same dimension. However, there is no natural way to choose an isomorphism between them, unless we pick some additional structure on V (such as a basis or a non-degenerate bilinear form).

On the other hand, we can construct an isomorphism ψ from V to $(V^*)^*$ as follows. If $v \in V$, then we define $\psi(v)$ to be the element of V^* given by

$$(\psi(v))(\alpha) = \alpha(v)$$

for all $\alpha \in V^*$. In other words, given a guy in V, we tell him to eat elements in V^* by allowing himself to be eaten.

Proposition 1.2. *ψ is an isomorphism.*

Proof. Since V and $(V^*)^*$ have the same dimension, it is enough to show that ψ is injective.

Suppose that $v \in V$, $v \neq 0$, and $\psi(v) = 0$. We wish to derive a contradiction.

Since $v \neq 0$, we can extend v to a basis $v_1 = v, v_2, \ldots, v_n$ for V. Then let α_1 defined as above. Then $\alpha_1(v) = 1 \neq 0$ and so we have a contradiction. Thus ψ is injective as desired. $\square$

From this proposition, we derive the following useful result.

Corollary 1.3. *Let $\alpha_1, \dots, \alpha_n$ be a basis for V^*. Then there exists a basis $v_1, \dots, v_n$ for V such that*

$$\alpha_i(v_j) = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$$

for all i, j.

Proof. Let $w_1, \dots, w_n$ be the dual basis to $\alpha_1, \dots, \alpha_n$ in $(V^*)^*$. Since ψ is invertible, ψ^{-1} exists. Define $v_i = \psi^{-1}(w_i)$. Since $w_1, \dots, w_n$ is a basis, so is $v_1, \dots, v_n$. Checking through the definitions shows that $v_1, \dots, v_n$ have the desired properties. □

2 Bilinear forms

Let V be a vector space. We denote the set of all bilinear forms on V by $(V^*)^{\otimes 2}$. We have already seen that this set is a vector space.

Similarly, we have the subspaces $Sym^2 V^*$ and $\Lambda^2 V^*$ of symmetric and skew-symmetric bilinear forms.

Proposition 2.1. $(V^*)^{\otimes 2} = Sym^2 V^* \oplus \Lambda^2 V^*$

Proof. Clearly, $Sym^2 V^* \cap \Lambda^2 V^* = 0$, so it suffices to show that any bilinear form is the sum of a symmetric and skew-symmetric bilinear form. Let H be a bilinear form. Let $\hat{H}$ be the bilinear form defined by

$$\hat{H}(v_1, v_2) = H(v_2, v_1)$$

Then $(H + \hat{H})/2$ is symmetric and $(H - \hat{H})/2$ is skew-symmetric. Hence $H = (H+\hat{H})/2 + (H-\hat{H})/2$ is the sum of a symmetric and skew-symmetric form. □

If $\alpha, \beta \in V^*$, then we can define a bilinear form $\alpha \otimes \beta$ as follows.

$$(\alpha \otimes \beta)(v_1, v_2) = \alpha(v_1)\beta(v_2)$$

for $v_1, v_2 \in V$.

We can also define a symmetric bilinear form $\alpha \cdot \beta$ by

$$(\alpha \cdot \beta)(v_1, v_2) = \alpha(v_1)\beta(v_2) + \alpha(v_2)\beta(v_1)$$

and a skew-symmetric bilinear from $\alpha \wedge \beta$ by

$$(\alpha \wedge \beta)(v_1, v_2) = \alpha(v_1)\beta(v_2) - \alpha(v_2)\beta(v_1)$$

These operations are linear in each variable. In other words

$$\alpha \otimes (\beta + \gamma) = \alpha \otimes \beta + \alpha \otimes \gamma$$

and similar for the other operations.

Example 2.2. Take $V = \mathbb{R}^2$. Let α_1, α_2 be the standard dual basis for V^*, so that

$$\alpha_1(x_1, x_2) = x_1, \;\; \alpha_2(x_1, x_2) = x_2$$

Then $\alpha_1 \otimes \alpha_2$ is given by

$$(\alpha_1 \otimes \alpha_2)((x_1, x_2), (y_1, y_2)) = x_1 y_2$$

Similarly $\alpha_1 \wedge \alpha_2$ is the standard symplectic form on $\mathbb{R}^2$, given by

$$(\alpha_1 \wedge \alpha_2)((x_1, x_2), (y_1, y_2)) = x_1 y_2 - x_2 y_1$$

$\alpha_1 \cdot \alpha_2$ is the symmetric bilinear form of signature $(1, 1)$ on $\mathbb{R}^2$ given by

$$(\alpha_1 \cdot \alpha_2)((x_1, x_2), (y_1, y_2)) = x_1 y_2 + x_2 y_1$$

The standard positive definite bilinear form on $\mathbb{R}^2$ (the dot product) is given by $\alpha_1 \cdot \alpha_1 + \alpha_2 \cdot \alpha_2$.

3 Multilinear forms

Let V be a vector space.

We can consider k-forms on V, which are maps

$$H : V \times \cdots \times V \to \mathbb{F}$$

which are linear in each argument. In other words

$$\begin{aligned} H(av_1, \ldots, v_k) &= aH(v_1, \ldots, v_k) \\ H(v + w, v_2, \ldots, v_k) &= H(v, v_2, \ldots, v_k) + H(w, v_2, \ldots, v_k) \end{aligned}$$

for $a \in \mathbb{F}$ and $v, w, v_1, \ldots, v_k \in V$, and similarly in all other arguments.

H is called symmetric if for each i, and all $v_1, \ldots, v_k$,

$$H(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, v_{i+2}, \ldots, v_n) = H(v_1, \ldots, v_{i-1}, v_{i+1}, v_i, v_{i+2}, \ldots, v_n)$$

H is called skew-symmetric (or alternating) if for each i, and all $v_1, \ldots, v_k$,

$$H(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, v_{i+2}, \ldots, v_n) = -H(v_1, \ldots, v_{i-1}, v_{i+1}, v_i, v_{i+2}, \ldots, v_n)$$

The vector space of all k-forms is denoted $(V^*)^{\otimes k}$ and the subspaces of symmetric and skew-symmetric forms are denote $Sym^k V^*$ and $\Lambda^k V^*$.

3.1 Permutations

Let S_k denote the set of bijections from $\{1, \ldots, k\}$ to itself (called a permutation). S_k is also called the symmetric group. It has $k!$ elements. The permutations occuring in the definition of symmetric and skew-symmetric forms are

called simple transpositions (they just swap i and $i+1$). Every permutation can be written as a composition of simple transpositions.

From this it immediately follows that if H is symmetric and if $\sigma \in S_k$, then

$$H(v_1, \ldots, v_k) = H(v_{\sigma(1)}, \cdots, v_{\sigma(n)})$$

There is a function $\varepsilon : S_k \to \{1, -1\}$ called the sign of a permutation, which is defined by the conditions that $\varepsilon(\sigma) = -1$ if σ is a simple transposition and

$$\varepsilon(\sigma_1\sigma_2) = \varepsilon(\sigma_1)\varepsilon(\sigma_2)$$

for all $\sigma_1, \sigma_2 \in S_k$.

The sign of a permutation gives us the behaviour of skew-symmetric k-forms under permuting the arguments. If H is skew-symmetric and if $\sigma \in S_k$, then

$$H(v_1, \ldots, v_k) = \varepsilon(\sigma) H(v_{\sigma(1)}, \cdots, v_{\sigma(n)})$$

3.2 Iterated tensors, dots, and wedges

If H is a $k-1$-form and $\alpha \in V^*$, then we define $H \otimes \alpha$ to be the k-form defined by

$$(H \otimes \alpha)(v_1, \ldots, v_k) = H(v_1, \ldots, v_{k-1})\alpha(v_k)$$

Similarly, if H is a symmetric $k-1$-form and $\alpha \in V^*$, then we define $H \cdot \alpha$ to be the k-form defined by

$$(H \otimes \alpha)(v_1, \ldots, v_k) = H(v_1, \ldots, v_{k-1})\alpha(v_k) + \cdots + H(v_2, \ldots, v_k)\alpha(v_1)$$

It is easy to see that $H \cdot \alpha$ is a symmetric k-form.

Similarly, if H is a skew-symmetric $k-1$-form and $\alpha \in V^*$, then we define $H \wedge \alpha$ to be the k-form defined by

$$(H \otimes \alpha)(v_1, \ldots, v_k) = H(v_1, \ldots, v_{k-1})\alpha(v_k) - \cdots \pm H(v_2, \ldots, v_k)\alpha(v_1)$$

It is easy to see that $H \wedge \alpha$ is a skew-symmetric k-form.

From these definitions, we see that if $\alpha_1, \ldots, \alpha_k \in V^*$, then we can iteratively define

$$\alpha_1 \otimes \cdots \otimes \alpha_k := ((\alpha_1 \otimes \alpha_2) \otimes \alpha_3) \otimes \cdots \otimes \alpha_k$$

and similar definitions for $\alpha_1 \cdots \alpha_k$ and $\alpha_1 \wedge \cdots \wedge \alpha_k$.

When we expand out the definitions of $\alpha_1 \cdots \alpha_k$ and $\alpha_1 \wedge \cdots \wedge \alpha_k$ there will be $k!$ terms, one for each element of S_k.

For any $\sigma \in S_k$, we have

$$\alpha_1 \cdots \alpha_k = \alpha_{\sigma(1)} \cdots \alpha_{\sigma(k)}$$

and

$$\alpha_1 \wedge \cdots \wedge \alpha_k = \varepsilon(\sigma)\alpha_{\sigma(1)} \wedge \cdots \wedge \alpha_{\sigma(k)}$$

The later property implies that $\alpha_1 \wedge \cdots \wedge \alpha_k = 0$ if $\alpha_i = \alpha_j$ for some $i \neq j$.

The following result is helpful in understanding these iterated wedges.

Theorem 3.1. *Let $\alpha_1, \dots, \alpha_k \in V^*$.*

$$\alpha_1 \wedge \cdots \wedge \alpha_k = 0 \text{ if and only if } \alpha_1, \dots, \alpha_k \text{ are linearly dependent}$$

Proof. Suppose that $\alpha_1, \dots, \alpha_k$ is linearly dependent. Then there exists $x_1, \dots, x_k$ such that

$$x_1\alpha_1 + \cdots + x_k\alpha_k = 0$$

and not all $x_1, \dots, x_k$ are zero. Assume that $x_k \neq 0$. Let $H = \alpha_1 \wedge \cdots \wedge \alpha_{k-1}$ and let us apply $H\wedge$ to both sides of this equation. Using the above results and the linearity of $\wedge$, we deduce that

$$x_k\alpha_1 \wedge \cdots \wedge \alpha_{k-1} \wedge \alpha_k = 0$$

which implies that $\alpha_1 \wedge \cdots \wedge \alpha_k = 0$ as desired.

For the converse, suppose that $\alpha_1, \dots, \alpha_k$ are linearly independent. Then we can extend $\alpha_1, \dots, \alpha_k$ to a basis $\alpha_1, \dots, \alpha_n$ for V^*. Let $v_1, \dots, v_n$ be the dual basis for V. Then

$$(\alpha_1 \wedge \cdots \wedge \alpha_k)(v_1, \dots, v_k) = 1$$

and so $\alpha_1 \wedge \cdots \wedge \alpha_k \neq 0$. $\square$

The same method of proof can be used to prove the following result.

Theorem 3.2. *Let $v_1, \dots, v_k \in V$. Then there exists $H \in \Lambda^k V^*$ such that $H(v_1, \dots, v_k) \neq 0$ if and only if $v_1, \dots, v_k$ are linearly independent.*

In particular this theorem shows that $\Lambda^k V^* = 0$ if $k > dim V$.

3.3 Bases and dimension

We will now describe bases for our vector spaces of k-forms.

Theorem 3.3. *Let $\alpha_1, \dots, \alpha_n$ be a basis for V^*.*

(i) $\{\alpha_{i_1} \otimes \cdots \otimes \alpha_{i_k}\}_{1 \leq i_1, \dots, i_k \leq n}$ *is a basis for* $(V^*)^{\otimes k}$.

(ii) $\{\alpha_{i_1} \cdots \alpha_{i_k}\}_{1 \leq i_1 \leq \cdots \leq i_k \leq n}$ *is a basis for* $Sym^k V^*$.

(iii) $\{\alpha_{i_1} \wedge \cdots \wedge \alpha_{i_k}\}_{1 \leq i_1 < \cdots < i_k \leq n}$ *is a basis for* $\Lambda^k V^*$.

Proof. We give the proof for the case of $(V^*)^{\otimes k}$ as the other cases are similar. So simplify the notation, let us assume that $k = 2$.

Let us first show that every bilinear form is a linear combination of $\{\alpha_i \otimes \alpha_j\}$. Let H be a bilinear form. Let $v_1, \dots, v_n$ be the basis of V dual to $\alpha_1, \dots, \alpha_n$. Let $c_{ij} = H(v_i, v_j)$ for each i, j. We claim that

$$H = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} \alpha_i \otimes \alpha_j$$

Since both sides are bilinear forms, it suffices to check that they agree on all pairs (v_k, v_l) of basis vectors. By definition $H(v_k, v_l) = c_{kl}$. On the other hand,

$$\left(\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\alpha_i \otimes \alpha_j\right)(v_k, v_l) = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\alpha_i(v_k)\alpha_j(v_l) = c_{kl}$$

and so the claim follows.

Now to see that $\{\alpha_i \otimes \alpha_j\}$ is a linearly independent set, just note that if

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\alpha_i \otimes \alpha_j = 0,$$

then by evaluating both sides on (v_i, v_j), we see that $c_{ij} = 0$ for all i, j. $\square$

Example 3.4. Take $n = 2, k = 2$. Then our bases are

$$\alpha_1 \otimes \alpha_1, \alpha_1 \otimes \alpha_2, \alpha_2 \otimes \alpha_1, \alpha_2 \otimes \alpha_2$$

and

$$\alpha_1 \cdot \alpha_1, \alpha_1 \cdot \alpha_2, \alpha_2 \cdot \alpha_2$$

and

$$\alpha_1 \wedge \alpha_2$$

Corollary 3.5. *The dimension of $(V^*)^{\otimes k}$ is n^k, the dimension of $Sym^k V^*$ is $\binom{n+k-1}{k}$ and the dimension of $\Lambda^k V^*$ is $\binom{n}{k}$.*

Tensor products

Joel Kamnitzer

April 5, 2011

1 The definition

Let V, W, X be three vector spaces. A *bilinear map* from $V \times W$ to X is a function $H : V \times W \to X$ such that

$$H(av_1 + v_2, w) = aH(v_1, w) + H(v_2, w) \text{ for } v_1, v_2 \in V, w \in W, a \in \mathbb{F}$$
$$H(v, aw_1 + w_2) = aH(v, w_1) + H(v, w_2) \text{ for } v \in V, w_1, w_2 \in W, a \in \mathbb{F}$$

Let V and W be vector spaces. A *tensor product* of V and W is a vector space $V \otimes W$ along with a bilinear map $\phi : V \times W \to V \otimes W$, such that for every vector space X and every bilinear map $H : V \times W \to X$, there exists a unique linear map $T : V \otimes W \to X$ such that $H = T \circ \phi$.

In other words, giving a linear map from $V \otimes W$ to X is the same thing as giving a bilinear map from $V \times W$ to X.

If $V \otimes W$ is a tensor product, then we write $v \otimes w := \phi(v \otimes w)$. Note that there are two pieces of data in a tensor product: a vector space $V \otimes W$ and a bilinear map $\phi : V \times W \to V \otimes W$.

Here are the main results about tensor products summarized in one theorem.

Theorem 1.1. *(i) Any two tensor products of V, W are isomorphic.*

(ii) V, W has a tensor product.

(iii) If $v_1, \dots, v_n$ is a basis for V and $w_1, \dots, w_m$ is a basis for W, then

$$\{v_i \otimes w_j\}_{1 \le i \le n, 1 \le j \le m}$$

is a basis for $V \otimes W$.

In particular, the last part of the theorem shows we can think of elements of $V \otimes W$ as $n \times m$ matrices with entries in $\mathbb{F}$.

2 Existence

We will start by proving that the tensor product exists. To do so, we will construct an explicit tensor product. This construction only works if V, W are finite-dimensional.

Let $B(V^*, W^*; \mathbb{F})$ be the vector space of bilinear maps $H : V^* \times W^* \to \mathbb{F}$. If $v \in V$ and $w \in W$, then we can define a bilinear map $v \otimes w$ by

$$(v \otimes w)(\alpha, \beta) = \alpha(v)\beta(w).$$

Just as we saw before, we have the following result.

Theorem 2.1. *Let $v_1, \ldots, v_n$ be a basis for V and let $w_1, \ldots, w_m$ be a basis for W. Then $\{v_i \otimes w_j\}_{1 \leq i \leq n, 1 \leq j \leq m}$ is a basis for $B(V^*, W^*; \mathbb{F})$*

Now, we define a map $\phi : V \times W \to B(V^*, W^*; \mathbb{F})$ by $\phi(v, w) = v \otimes w$.

Theorem 2.2. *$B(V^*, W^*; \mathbb{F})$ along with ϕ is a tensor product for V, W.*

Note that this proves parts (ii) and (iii) of our main theorem.

Proof. Fix bases $v_1, \ldots, v_n$ for V and $w_1, \ldots, w_m$ for W.

Let X be a vector space and let $H : V \times W \to X$ be a bilinear map. We define a linear map $T : V \otimes W \to X$ by defining it on our basis as $T(v_i \otimes w_j) = H(v_i, w_j)$. Then $T \circ \phi$ and H are two bilinear maps from $V \times W$ to X which agree on basis vectors, hence they are equal. (Note that it is easy to show that for any $(v, w) \in V \times W$, $T(v \otimes w) = H(v, w)$.)

Finally, note that T is the unique linear map with this property, since it is determined on the basis for $B(V^*, W^*, \mathbb{F})$. □

Using the same ideas, it is easy to see that $L(V^*, W)$, $L(W^*, V)$, and $B(V, W; \mathbb{F})^*$ are all also tensor products of V, W.

3 Uniqueness

Now we prove uniqueness. Here is the precise statement.

Theorem 3.1. *Let $(V \otimes W)_1, \phi_1$ and $(V \otimes W)_2, \phi_2$ be two tensor products of V, W. Then there exists a unique isomorphism $T : (V \otimes W)_1 \to (V \otimes W)_2$ such that $\phi_2 = T \circ \phi_1$.*

Proof. Let us apply the definition of tensor product to $(V \otimes W)_1, \phi_1$ with the role of X, H taken by $(V \otimes W)_2, \phi_2$. By the definition, we obtain a (unique) linear map $T : (V \otimes W)_1 \to (V \otimes W)_2$ such that $\phi_2 = T \circ \phi_1$.

Reversing the roles of $(V \otimes W)_1, \phi_1$ and $(V \otimes W)_2, \phi_2$, we find a linear map $S : (V \otimes W)_2 \to (V \otimes W)_1$ such that $\phi_1 = S \circ \phi_2$.

We claim that $T \circ S = I_{(V \otimes W)_2}$ and $S \circ T = I_{(V \otimes W)_1}$ and hence T is an isomorphism. We will now prove $S \circ T = I_{(V \otimes W)_1}$.

Note that $(S \circ T) \circ \phi_1 = S \circ \phi_2 = \phi_1$ by the above equations. Now, apply the definition of tensor product to $(V \otimes W)_1, \phi_1$ with the role of X, H taken by $(V \otimes W)_1, \phi_1$. Then both $S \circ T$ and $I_{(V \otimes W)_1}$ can play the role of T. So by the uniqueness of "T" in the definition, we conclude that $S \circ T = I_{(V \otimes W)_1}$ as desired.

Finally to see that the T that appears in the statement of the theorem is unique, we just note from the first paragraph of this proof, it follows that there is only one linear map T such that $\phi_2 = T \circ \phi_1$. $\square$